



OICE

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/10/028,748

DATE: 02/06/2002  
 TIME: 09:38:39

Input Set : N:\Crf3\RULE60\10028748.raw  
 Output Set: N:\CRF3\02062002\J028748.raw

ENTERED

## SEQUENCE LISTING

## 1 (1) GENERAL INFORMATION:

- 2 (i) APPLICANT: Mack, David H.  
 3 (ii) TITLE OF INVENTION: COMPUTER-AIDED VISUALIZATION OF  
 4 EXPRESSION COMPARISON  
 5 (iii) NUMBER OF SEQUENCES: 2  
 6 (iv) CORRESPONDENCE ADDRESS:  
 7 (A) ADDRESSEE: Joe Liebeschuetz of Townsend and Townsend and  
 8 Crew LLP  
 9 (B) STREET: Two Embarcadero Center, Eighth Floor  
 10 (C) CITY: San Francisco  
 11 (D) STATE: CA  
 12 (E) COUNTRY: USA  
 13 (F) ZIP: 94111-3834  
 14 (v) COMPUTER READABLE FORM:  
 15 (A) MEDIUM TYPE: Floppy disk  
 16 (B) COMPUTER: IBM PC compatible  
 17 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 18 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
 22 (vi) CURRENT APPLICATION DATA:  
 C--> 23 (A) APPLICATION NUMBER: US/10/028,748  
 C--> 24 (B) FILING DATE: 21-Dec-2001  
 25 (C) CLASSIFICATION:  
 28 (vii) PRIOR APPLICATION DATA:  
 29 (A) APPLICATION NUMBER: US/09/020,743  
 30 (B) FILING DATE: 09-Feb-1998  
 34 (viii) ATTORNEY/AGENT INFORMATION:  
 35 (A) NAME: Liebeschuetz, Joe  
 36 (B) REGISTRATION NUMBER: 37,505  
 37 (C) REFERENCE/DOCKET NUMBER: 018547034800US  
 38 (ix) TELECOMMUNICATION INFORMATION:  
 39 (A) TELEPHONE: (650) 326-2400  
 40 (B) TELEFAX: (650) 326-2422  
 41 (2) INFORMATION FOR SEQ ID NO: 1:  
 42 (i) SEQUENCE CHARACTERISTICS:  
 43 (A) LENGTH: 2691 base pairs  
 44 (B) TYPE: nucleic acid  
 45 (C) STRANDEDNESS: unknown  
 W--> 46 (D) TOPOLOGY: not relevant  
 47 (ii) MOLECULE TYPE: DNA (genomic)  
 48 (vi) ORIGINAL SOURCE:  
 49 (A) ORGANISM: Homo sapiens  
 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,748

DATE: 02/06/2002

TIME: 09:38:39

Input Set : N:\Crif3\RULE60\10028748.raw

Output Set: N:\CRF3\02062002\J028748.raw

51	GGAGACAGAC	AGACAGCTGG	CAAGAGGCAG	CCTGGGGGCC	ACAGCTGCTT	CAGCAGACCT	60
52	CATGGCTGAG	TGAGCCTCCC	CTGGGCCCAG	CACCCACCT	CAGCATGGTC	CAAGCCCATG	120
53	GGGGGCGCTC	CAGAGCACAG	CCGTTGACCT	TGTCTTGGG	GGCAGCCATG	ACCCAGCCTC	180
54	CGCCTGAAAA	AACGCCAGCC	AAGAAGCATG	TGCGACTGCA	GGAGAGGCGG	GGCTCCAATG	240
55	TGGTCTTGAT	GCTGGACGTT	CGGTCCCTGG	GGGCCGTAGA	ACCCATCTGC	TCTGTGAACA	300
56	CACCCCGGGA	GGTCACCCTA	CACTTTCTGC	GCACTGCTGG	ACACCCCTT	ACCCGCTGGG	360
57	CCCTTCAGCG	CCAGCCACCC	AGCCCCAAGC	AACTGGAAGA	AGAATTCTTG	AAGATCCCTT	420
58	CAAACCTTGT	CAGCCCCGAA	GACCTGGACA	TCCCTGGCCA	CGCCTCCAAG	GACCGATACA	480
59	AGACCATCTT	GCCAAATCCC	CAGAGCCGTG	TCTGTCTAGG	CCGGGCACAG	AGCCAGGAGG	540
60	ACGGAGATTA	CATCAATGCC	AACTACATCC	GAGGCTATGA	CGGGAAGGAG	AAGGTCTACA	600
61	TTGCCACCCA	GGGCCCCATG	CCCAACACTG	TGTCGGACTT	CTGGGAGATG	GTGTGGCAAG	660
62	AGGAAGTGTC	CCTCATTGTC	ATGCTCACTC	AGCTCCGAGA	GGGCAAGGAG	AAATGTGTCC	720
63	ACTACTGGCC	CACAGAAGAG	GAAACCTATG	GACCCTTCCA	GATCCGCATC	CAGGACATGA	780
64	AAGAGTGCCC	AGAATACACT	GTGCGGCAGC	TCACCATCCA	GTACCAGGAA	GAGCGCCGGT	840
65	CAGTAAAGCA	CATCCTCTTT	TCGGCCTGGC	CAGACCATCA	GACACCAGAA	TCAGCTGGGC	900
66	CCCTGCTGCG	CCTAGTGGCA	GAGGTGGAGG	AGAGCCCGGA	GACAGCCGCC	CACCCCGGGC	960
67	CTATCGTAGT	CCACTGCAGT	GCAGGGATTG	GCCGGACGGG	CTGCTTCATC	GCCACGCGAA	1020
68	TTGGCTGTCA	ACAGCTGAAA	CCCCGAGGAG	AAGTGGACAT	TCTGGGTATT	GTGTGCCAAC	1080
69	TGCGGCTAGA	CAGAGGGGGG	ATGATCCAGA	CGGACGAGCA	GTACCAGTTC	CTGCACCACA	1140
70	CTTTGGCCCT	GTATGCAGGC	CAGCTGCCTG	AGGAACCCAG	CCCCTGACCC	CTGCCACCCT	1200
71	CCGGTGGCCC	AGGTGCCTAC	CTCCCTCAAG	CCTGGGAAGT	CACAGGAAGC	AGCAGCAGTA	1260
72	AGGACAAGGG	GCCGGATTCC	AGGTCTTCAA	CACTGGCCAC	TCCTCTGCTT	CCTCTGTTGG	1320
73	CCCCAGATGG	ACAGTAAGGG	GAACCTCCAA	TGTCTCTCTG	AACTTAAAGA	CAGGAGCTGG	1380
74	CATTTATGAC	AGACAAAGAA	AGAAGCCCAG	GTGTCCTGGT	GTTCTCTGAG	ACACTCTTTG	1440
75	TGAGCTTCAG	TTTCCTGTTC	TATAACATGA	ACATAAGTGC	TTAGCTGCCA	TGAGGGAAAA	1500
76	GTAATGAGAG	AAGTTTCTAG	AAGCCACTCC	AGCCACTCCT	TCCTGGGGCT	GACAAAAGGG	1560
77	TGATTCCAAG	ATCATCCTTC	ACCCGAGGTC	CTGCCCAAGC	ACAGGCCAGA	TGCAAGAAATG	1620
78	GGGAAAAGTC	TGGTCCTGAT	CTCCAAGTCT	CAACATCCTA	TCAGTGACTC	TGCTCCCTGA	1680
79	CCACACATCG	GAAGGGCTGG	ATGACCCCAA	TCAAAAGAAA	GAACAAGGAC	TCTGGTTACC	1740
80	CTTGCCCTCC	ACCCATGTGT	CATAAGAGTA	GGCTACAGAG	GTGACCAGGC	CTGGCAGTTG	1800
81	AAATCTCTGG	AAGAGGGAAC	ATGTGGGGAC	TACTCAGAGG	CAAAGAGGAG	CTGCTCCTGC	1860
82	CTCCATGGTT	GCTGGCCTCT	CCCAACAACT	ACTCTTAGGG	AGGCTAAGCA	GTCTCTGTTT	1920
83	TGCTTCCATG	GCTCAAATAA	TACCCTGGGT	ATGCAGGACC	CACTATACCT	TGCATTTGCT	1980
84	GGTACACCTA	GAGAGCTTGG	CTGTTTCCAA	AAACAATCAG	GGTCATAACC	ATCCATGCAG	2040
85	ACATGGAGGC	TCGGCTGAAC	CAGGACTCCT	CACTGTCTAC	CTGAGAGAAT	GAGCACCCCT	2100
86	CATCCATCTC	AGCATCAACA	CAATTTCCAG	GGGACCTCAG	GTCTACCTCA	GGACTGAACG	2160
87	CCACACCTCA	GGATTCTCTC	TCCTTGAATC	TGAGACTGGC	TGCCCATTCT	GAGATGGGGA	2220
88	TGAAGGTAAG	ATGCCGCATC	ACCAGGCACG	CCGCCCCCTGA	CAGCTGCCTT	GATACCAGCT	2280
89	CTCTGTGGAA	ACCCCGGAGG	AGTTGGATCT	GGAGAACAGC	TGGGCCTCCT	CACTCAGGAC	2340
90	TTCTCTCCTG	AAGAACACGC	AGTGCTAAAA	CTGAGGATGA	TTTCCCTAAT	GCTTCTGCTT	2400
91	GGCCTTATGG	AGGAGCTGCT	CCTTCCTTAC	AGCCTTGGGG	ATGGACTTGC	CCACACCTCC	2460
92	ACCTCCCCTG	AGCCCTGTGA	GAGGCACGAC	TGTCTATGCC	AATGAGGCTC	GGTGGGGGGC	2520
93	TCTCAAGTGC	CTGATCCTGC	CCTGGGCTCA	GAGCCAGCCC	AGAGGGAAGC	AACTGCACAG	2580
94	CCCCACAGGC	CCTCCCTGGC	ACTGTCCCCC	CAACCCCATC	TCAGAGCTCA	GAGGGTACAA	2640
95	GCTCCAGAAC	AGTAACCAAG	TGGGAAAATA	AAGACTTCTT	GGATGACTGA	C	2691
97	(2)	INFORMATION FOR SEQ ID NO: 2:					
98		(i) SEQUENCE CHARACTERISTICS:					
99		(A) LENGTH: 360 amino acids					
100		(B) TYPE: amino acid					

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,748

DATE: 02/06/2002

TIME: 09:38:39

Input Set : N:\Crf3\RULE60\10028748.raw

Output Set: N:\CRF3\02062002\J028748.raw

101 (C) STRANDEDNESS: not relevant  
 W--> 102 (D) TOPOLOGY: not relevant  
 103 (ii) MOLECULE TYPE: protein  
 104 (vi) ORIGINAL SOURCE:  
 105 (A) ORGANISM: Homo sapiens  
 106 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 107 Met Val Gln Ala His Gly Gly Arg Ser Arg Ala Gln Pro Leu Thr Leu  
 108 1 5 10 15  
 109 Ser Leu Gly Ala Ala Met Thr Gln Pro Pro Pro Glu Lys Thr Pro Ala  
 110 20 25 30  
 111 Lys Lys His Val Arg Leu Gln Glu Arg Arg Gly Ser Asn Val Ala Leu  
 112 35 40 45  
 113 Met Leu Asp Val Arg Ser Leu Gly Ala Val Glu Pro Ile Cys Ser Val  
 114 50 55 60  
 115 Asn Thr Pro Arg Glu Val Thr Leu His Phe Leu Arg Thr Ala Gly His  
 116 65 70 75 80  
 117 Pro Leu Thr Arg Trp Ala Leu Gln Arg Gln Pro Pro Ser Pro Lys Gln  
 118 85 90 95  
 119 Leu Glu Glu Glu Phe Leu Lys Ile Pro Ser Asn Phe Val Ser Pro Glu  
 120 100 105 110  
 121 Asp Leu Asp Ile Pro Gly His Ala Ser Lys Asp Arg Tyr Lys Thr Ile  
 122 115 120 125  
 123 Leu Pro Asn Pro Gln Ser Arg Val Cys Leu Gly Arg Ala Gln Ser Gln  
 124 130 135 140  
 125 Glu Asp Gly Asp Tyr Ile Asn Ala Asn Tyr Ile Arg Gly Tyr Asp Gly  
 126 145 150 155 160  
 127 Lys Glu Lys Val Tyr Ile Ala Thr Gln Gly Pro Met Pro Asn Thr Val  
 128 165 170 175  
 129 Ser Asp Phe Trp Glu Met Val Trp Gln Glu Glu Val Ser Leu Ile Val  
 130 180 185 190  
 131 Met Leu Thr Gln Leu Arg Glu Gly Lys Glu Lys Cys Val His Tyr Trp  
 132 195 200 205  
 133 Pro Thr Glu Glu Glu Thr Tyr Gly Pro Phe Gln Ile Arg Ile Gln Asp  
 134 210 215 220  
 135 Met Lys Glu Cys Pro Glu Tyr Thr Val Arg Gln Leu Thr Ile Gln Tyr  
 136 225 230 235 240  
 137 Gln Glu Glu Arg Arg Ser Val Lys His Ile Leu Phe Ser Ala Trp Pro  
 138 245 250 255  
 139 Asp His Gln Thr Pro Glu Ser Ala Gly Pro Leu Leu Arg Leu Val Ala  
 140 260 265 270  
 141 Glu Val Glu Glu Ser Pro Glu Thr Ala Ala His Pro Gly Pro Ile Val  
 142 275 280 285  
 143 Val His Cys Ser Ala Gly Ile Gly Arg Thr Gly Cys Phe Ile Ala Thr  
 144 290 295 300  
 145 Arg Ile Gly Cys Gln Gln Leu Lys Ala Arg Gly Glu Val Asp Ile Leu  
 146 305 310 315 320  
 147 Gly Ile Val Cys Gln Leu Arg Leu Asp Arg Gly Gly Met Ile Gln Thr  
 148 325 330 335  
 149 Asp Glu Gln Tyr Gln Phe Leu His His Thr Leu Ala Leu Tyr Ala Gly

## RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/10/028,748

TIME: 09:38:39

Input Set : N:\Crf3\RULE60\10028748.raw

Output Set: N:\CRF3\02062002\J028748.raw

150				340			345		350
151	Gln	Leu	Pro	Glu	Glu	Pro	Ser	Pro	
152			355				360		

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/028,748

DATE: 02/06/2002

TIME: 09:38:40

Input Set : N:\Crf3\RULE60\10028748.raw

Output Set: N:\CRF3\02062002\J028748.raw

L:23 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:24 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:46 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=1  
L:102 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=2